

## Mater Thesis/Internship position

Laboratoire d'Informatique  
Fondamentale et Appliquée  
64 avenue Jean Portalis  
37200 TOURS  
lifat.univ-tours.fr

Équipe Reconnaissance des  
Formes et Analyse d'Images  
www.rfai.li.univ-tours.fr

### Machine learning for time series prediction in environmental sciences

**Profile:** Academic level equivalent to a Master 2 in progress or Engineer in its 5<sup>th</sup> year in computer science

**Duration:** 4 to 6 months, from March-April 2024

**Affiliation:** Computer Science Lab of Université de Tours ([LIFAT](#)), Pattern Recognition and Image Analysis Group ([RFAI](#))

**Salary:** according to French legislation (around 600€/month in average); plus indemnity for local transports

**Supervisor:** Nicolas RAGOT (LIFAT-Tours)

**Skills:**

- a good experience in data analysis and machine learning (in python) is required
- some knowledge and experiences in deep learning and associated tools will be highly considered
- some knowledge in time series analysis and forecasting will be highly considered
- curiosity and ability to communicate and share your progress and to make written reports
- ability to propose solutions
- autonomy and good organization skills

**How to candidate:**

Send the following documents by e-mail to [nicolas.ragot \[at\] univ-tours.fr](mailto:nicolas.ragot[at]univ-tours.fr) before 1<sup>st</sup> of March 2024: a CV, motivation letter, your scores including bachelor degree, and references from teachers or people you worked with.

**Context:**

The JUNON project, driven by the BRGM, is granted from the Centre-Val de Loire region through ARD program (« Ambition Recherche Développement ») which goal is to develop a research & innovation pole around environmental resources (agriculture, forest, waters...). The main goal of JUNON is to elaborate digital services through large scale digital twins in order to improve the monitoring, understanding and prediction of environmental resources evolution and phenomena, for a better management of natural resources. Digital twins will allow to virtually reproduce natural processes and phenomena using combination of AI and environmental tools.

JUNON will focus on the elaboration of digital twins concerning quality and quantity of ground waters, as well as emissions of greenhouse gases and pollutants with health effects, at the scale of geographical area corresponding to the North part of the Centre-Val-de-Loire region.

Nicolas RAGOT  
+33 2 47 36 14 31  
[nicolas.ragot@univ-tours.fr](mailto:nicolas.ragot@univ-tours.fr)

UNIVERSITÉ DE TOURS

École polytechnique  
Département Informatique  
64 avenue Jean Portalis  
37200 TOURS

[polytech.univ-tours.fr](http://polytech.univ-tours.fr)

Laboratoire d'Informatique  
Fondamentale et Appliquée  
64 avenue Jean Portalis  
37200 TOURS  
lifat.univ-tours.fr

Équipe Reconnaissance des  
Formes et Analyse d'Images  
www.rfai.li.univ-tours.fr

These digital twins will rely on geological and meteorological knowledge and data (time series), as well as physic-based models.

The project actors are: BRGM, Université d'Orléans, Université de Tours, CNRS, INRAE, and ATOS and ANTEA companies. There are 5 work packages (WP):

1. User's needs and geological knowledge for ground water
2. User's needs and biological/chemical knowledge about pollutants and greenhouse gases
3. Data management and data mining
4. Time series prediction
5. Aggregation and realization of digital twins themselves

The Master Thesis/internship position will be in the WP 4, focused on the prediction of pollutants in the air. There will be strong interactions with WP2 and 3. The work will be supervised by the LIFAT - RFAI and you will have to interact with PhD student and postdoc in WP4 working on prediction on groundwater level.

#### Goals:

The goal of this internship will be to analyze data and to build prediction models able using meteorological data as well as measurements of pollutants observed in the past (other factors could also be included). To help in this task, the candidate will have the opportunity to study and get inspiration from models that have are currently tested for the prediction of ground water level.

#### Methodology:

- 1 A state of the art will be made with emphasizes on pollutants and greenhouse gases level prediction. A comparison with other environmental time series prediction will be made, especially considering prediction of groundwater levels.
- 2 Data cleaning and preparation will be done, based on known benchmarks as well as experts knowledge from WP2.
- 3 Defining an experimental protocol.
- 4 Experimenting classical models from state of the art.
- 5 Implementing and adapting a deep learning model inspired from work done on ground water level prediction.
- 6 Evaluation and comparison of the prediction models based on several criteria (performances at short, mid-term; number of data needed for learning/adapting the model; ability to transfer the model for data observed from another location...)
- 7 Improvements of the models.
- 8 Documentation, reports and cleaning of the code to make it reusable (using Git)

Nicolas RAGOT  
+33 2 47 36 14 31  
nicolas.ragot@univ-tours.fr

UNIVERSITÉ DE TOURS

École polytechnique  
Département Informatique  
64 avenue Jean Portalis  
37200 TOURS

polytech.univ-tours.fr

Laboratoire d'Informatique  
Fondamentale et Appliquée  
64 avenue Jean Portalis  
37200 TOURS  
lifat.univ-tours.fr

Équipe Reconnaissance des  
Formes et Analyse d'Images  
www.rfai.li.univ-tours.fr

### Hosting group:

The [RFAI group](#) (Pattern Recognition and Image Analysis) is part of the [LIFAT](#) (EA 6300) computer science lab. The group is composed of 5 Professors, 3 HDR (associate professors habilitated), 10 associate professors, 9 PhDs plus 4 co-supervised PhDs in other universities. The group is working mainly on pattern recognition and machine learning for image/video analysis and temporal data with application domains mainly in health, environment and humanities. The group has access to several computing resources and especially to [Leto](#) computer (CPU nodes + 4 gpu nodes with 4 Nvidia Tesla v100 each).

### Bibliography:

- S. Li, X. Jin, Y. Xuan, X. Zhou, W. Chen, Y.-X. Wang, and X. Yan. *Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting*. In Advances in Neural Information Processing Systems, pp. 5243–5253, 2019
- G. Zerveas, S. Jayaraman, D. Patel, A. Bhamidipaty, C. Eickhoff. *A Transformer-based Framework for Multivariate Time Series Representation Learning*. arXiv:2010.02803
- J. Ma, Zheng Shou, Alireza Zareian, Hassan Mansour, A. Vetro, and S. Chang. Cds: Cross-dimensional self-attention for multivariate, geo-tagged time series imputation. ArXiv, abs/1905.09904, 2019
- Sijie Yan and Yuanjun Xiong and Dahua Lin, *Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition*, AAAI'18, arXiv:1801.07455, [paper with code](#), 2018

Nicolas RAGOT  
☎ +33 2 47 36 14 31  
✉ nicolas.ragot@univ-tours.fr

### UNIVERSITÉ DE TOURS

📍 École polytechnique  
Département Informatique  
64 avenue Jean Portalis  
37200 TOURS  
🌐 polytech.univ-tours.fr